Evaluating the Consequences of Zero-Rating: Guidance for Regulators and Adjudicators

Bronwyn Howell & Roslyn Layton

August 2016

Abstract

In February 2015, the Federal Communications Commission's Open Internet Order reclassified broadband Internet access service (BIAS) under Title II of the Communications Act. The Order calls for case-by-case analysis by the Commission of alleged breaches of a general conduct standard in respect of usage-based pricing and "zero-rating" in agreements between BIAS operators and end consumers. These provisions enable to Commission to investigate cases such as T-Mobile's zero-rating of traffic to and from its Binge On video streaming platform.

The granting of discretionary powers of inquiry on a case-by-case basis raises the question of what evidence of real or potential harm would be sufficient to justify the commitment of scarce regulatory or judicial resources to the investigation of any specific example? The question is not limited to the US context. Similar provisions in exist in the European Commission's Net Neutrality Regulation. Furthermore, the answers are potentially important for competition law authorities in assessing cases of alleged harm and potential future abuses that might become more relevant as merger activity between BIAS operators and applications providers, specifically Content Distribution Networks (CDNs), increases.

This paper contributes towards the formulation of a principled economics-based approach under which to assess the potential harms and enhancements to be considered in evaluating the economic effects of any given example of zero-rating. Its primary contribution is five questions, derived from a combined economic and strategic analysis of the potential trade-offs that emerge from theoretical modelling of interactions in a complex internet ecosystem characterised by multiple two-sided platforms. Asking these questions will assist regulators and adjudicators to identify key elements in the case facts that indicate the need for different trade-offs and strategic motivations to be taken into account when making decisions. The questions posed here augment existing guidelines, such as BEREC's (2016) advice to regulators, which focus predominantly upon compliance with legal provisions rather than facilitating analysis if the economic trade-offs involved in any individual case of zero-rating.

1. Introduction

In February 2015, the FCC's Open Internet Order reclassified broadband Internet access service (BIAS) under Title II of the Communications Act. The Order enacted three bright-line rules: no blocking; no throttling (impairing or degrading content); and a per se ban on "paid prioritization" – the favouring of some traffic over other traffic, such as offering higher speeds or quality of service in exchange for third-party payment or to benefit an affiliated entity. Notably, three significant exceptions were made to the application of the bright line Open Internet rules. First, they do not apply to interconnection agreements between BIAS providers and edge providers, content delivery networks (CDNs) or backbone networks. Furthermore, two other practices – usage-based pricing (notably the use of data caps) at the consumer level and "zero-rating" (exemptions by broadband providers of certain edge providers or groups of providers form data caps) – were not specifically banned. Instead, they are subject to a general conduct standard: the Order calls for case-by-case analysis of complaints made to the Commission about these three practices (Bring, *et al*., 2015).

The Order represents a pragmatic compromise by the FCC to the bitterly-divided factions of the Network Neutrality debate. The bright-line rules prohibiting blocking, throttling and paid prioritization prevent network operators physically discriminating between packets in transit, so effectively impose "net neutrality as a non-discrimination rule" (Schuett, 2010 p 1). This supports the view of pro-neutrality advocates (e.g. Lee & Wu, 2009; Van Schewick, 2015) that in the physical movement of data, the Internet should remain "open, nondiscriminatory and largely managed by users rather than carriers" (Frieden, 2008 p 5).

The exceptions, however, acknowledge the complex economic trade-offs cited by net neutrality objectors. These trade-offs take place within an Internet ecosystem consisting of multiple intertwined two-sided platforms, of which BIAS operators are just one of many in the commercial chain linking senders and receivers of data (see, for example, Faulhaber, 2011; Krämer *et al.*, 2013; Greenstein, *et al*., 2016). By forbearing from imposing "net neutrality as a zero-price rule" (Schuett, 2010 p 1,) the FCC has conceded that the economic tradeoffs are extremely complex, and understanding of how they will play out in both the short and long-run is currently both new and imperfect (Greenstein, *et al.*, 2016). Hence simplistic calls for wide-reaching rules imposing strict neutrality or refraining entirely from intervening are unhelpful for policy formulation and rule-making. A measured case-by-case approach to evaluating specific instances of usage-based pricing and zero-rating facilitates growth in understanding, at the same time as it mitigates the risk of closing down potentially welfare-enhancing pricing options in a highly dynamic and rapidly-evolving Internet ecosystem.

1.1 Identifying Relevant Cases For Inquiry

Nonetheless, opting for case-by-case consideration raises the question of how the FCC (or other regulatory and adjudication bodies with similar regulatory and legislative approaches) will identify the specific instances of usage-based pricing or zero-rating expected to be potentially so detrimental to total welfare that they justify committing the non-trivial sums required for a full investigation. It also raises the question of the appropriate forum in which to undertake any inquiry. Whilst the FCC is specifically interested in the activities of Internet Service Providers (ISPs) providing the 'last mile' of Internet connectivity to end consumers, their activities do not occur in isolation. Rather, they form part of the complex interactions between multiple participants within the wider Internet ecosystem.

Within the Internet ecosystem, the boundaries between ISPs' roles as providers of last mile data transport and content providers are already blurred for cable providers, and are increasingly challenged by ISPs establishing proprietary content distribution networks (CDNs) (e.g. T-Mobile and Binge-On), proposed mergers between ISPs and major broadcasters (e.g. Vodafone and SKY Television in New Zealand) and co-ordinated actions between ISPs and CDNs (Kourandi, *et al.,* 2015; Greenstein, *et al.,* 2016). The complexity of interactions is amplified by the reality that many of the participants in different parts of the ecosystem operate as two- (or multi-) sided platforms, and in many cases are interacting with the same stakeholders. Systemic interrelationships mean that decisions made by one platform about the optimal way in which it facilitates interaction between two sets of stakeholders in respect one functional area of the ecosystem may influence the interactions between the same stakeholders on another platform in another functional area. Neither is it straightforward when one stakeholder can appear in two different capacities on either side of a single platform.

This suggests that it may be more appropriate to assess zero-rating cases under the more general provisions of Competition Law (Antitrust) where a broader view of market and ecosystem activity is possible, than under the more restricted scope of a regulatory agency focused on only one subset of market interactions and a subset of participants. This approach enables a broader consideration of the economic implications of the relevant pricing structures than may be possible when assessing them against compliance with a narrower set of regulatory rules.

However, neither is a competition law-based approach unproblematic. The complexities due to the interactions of common stakeholders in multiple two-sided platforms pose challenges to deciding upon on the relevant market definitions for any inquiry (Filistrucchi *et al.,* 2014; Evans & Schmalensee, 2012). Furthermore, even if agreement can be reached on the relevant markets for analysis, as extant market power is not a necessary condition for offering plans with either data caps or zero-rated usage, it is not axiomatic that their adoption,

even by a party with market power, is intended to foreclose competition in either the ISP or CDN markets. Nor is their adoption necessarily anti-competitive or harmful to total welfare. For example, Howell (2010) demonstrates that metered ISP tariffs – of which plans with data caps are one variant (Economides & Hermalin, 2015) - are a logical strategic response in a maturing Internet access market with heterogeneous consumers exhibiting an asymmetric distribution of data volume usage and the potential for ISP competition, where the initial offers in the nascent market were unmetered. Indeed, in this situation unmetered tariffs (where all traffic is effectively zero-rated) are sustainable only if all ISPs act jointly by agreeing not to offer a metered tariff. Regulatory prohibition of metered tariffs offers a low-cost means of constraining defection at the same time as it forecloses entry by ISP rivals seeking to take advantage of the price arbitrage opportunity offered by the skewed demand distribution. And whilst metered charging may reduce the volume of traffic generated by high-volume users, it is not a simple matter to empirically assess the effects on either advertising markets or the responses to incentives to create more applications and content as a consequence of the reduced volumes of traffic from a small number of users, as opposed to the increase in the number of subscribers with expected low usage volumes as the price of accessing an internet connection reduces, especially as the relevant markets for advertising and content creation may span many different geographical locations and competition authority jurisdictions.

The dilemma for regulators and policy-makers is therefore deciding what to take into account when developing a framework for assessing cases of zero-rating, and deciding how and when to intervene. Competition law is a helpful precedent so long as long-term interests of consumers is the relevant criterion for assessment, and the promotion of competition in discrete markets is a close proxy for those interests. However, when the interconnectedness of two-sided markets makes the assessment of those interests problematic, simplistic rules of thumb are less helpful. The long-term interests of consumers may not necessarily be the best proxy for the long-term interests of the ecosystem in all cases. Considering on a case-by-case basis the arrangements that maximise long-term total welfare appears to be the only viable value-neutral criterion to use for such analysis.

The dynamic nature of the total welfare criterion inevitably necessitates in some instances that the short-term interests of some ecosystem stakeholders may have to be compromised to achieve the long-term welfare maximum. Furthermore, what is optimal at one time of the ecosystem lifecycle may not be optimal at another (Yoo, 2010; Howell, 2014). To the extent that social preferences may override the economic considerations, then the costs of imposing those preferences must be recognised in both the total welfare foregone, and the transfers that those preferences engender between ecosystem stakeholders – noting also that a single stakeholder may participate in the ecosystem in multiple capacities, and these may vary

over the ecosystem lifecycle. Developing a framework for evaluating each case is made even more challenging by inherent uncertainties about how the ecosystem may evolve in the future. Decisions to intervene – or not – inevitably affect this evolution by creating or ruling out paths that the evolution may follow, making policy-makers and regulators stakeholders in the ecosystem.

The challenge is to develop guidelines for developing guidelines for evaluating cases and making decisions that facilitate the pursuit of the long-term welfare objective in a transparent and neutral manner.

1.2 The European (BEREC) Guidelines

The FCC is not alone in facing this challenge. The Body of European Regulators for Electronic Communications (BEREC) has produced a set of guidelines for member state regulators to assist in implementing the European Net Neutrality Regulation (BEREC, 2016).

The BEREC guidelines are derived from the pre-eminent objective of European telecommunications regulation to promote competition. Competition is proxied for long term benefit of consumers, and the end objective is to ultimately withdraw regulations and rely instead on competition law for mediating interaction within the internet ecosystem. Theoretically, at least, these guidelines should be underpinned by principles promoting welfare-enhancing competition where this is feasible, and intervening only when specific activities are proven to be harmful to long-term economic and social welfare.

In evaluating examples of zero-rating, BEREC advises European regulators to take account of (BEREC, 2016):

- whether the practices circumvent the general aims of the Regulation (to *"safeguard equal and non-discriminatory treatment of traffic"* and to "*guarantee the continued functioning of the internet ecosystem as an engine of innovation*");
- the market positions of the ISPs and CAPs involved;
- any effects on end-user rights of consumer and business end-users, e.g. reductions in the range of applications available, incentives for end-users to use certain applications, or whether there is a material reduction in end-user choice;
- any effects on end-user rights of Content and Application Providers (CAP), e.g. whether there is an effect on the range of content and applications which CAPs can provide, or whether they are materially discouraged from entering the market;
- the scale of the practice (e.g. the number of end-users subscribing to such an offer) and the extent to which end-users have access to alternative offers and / or other ISPs; and
- any effects on freedom of expression and media pluralism.

Although providing a handy checklist of criteria to consider when evaluating cases of zero-rating, the BEREC guidelines focus more upon ensuring legal compliance with the Net

Neutrality Regulation than to assessing their potential or actual economic effects in the internet ecosystem (Layton, 2016). For example, nowhere in the guidelines is it identified that the Regulation concerns markets where multiple two-sided platforms are interacting. The emphasis on legal compliance reflects the scope of the regulation, which established "common rules to safeguard the non-discriminatory treatment of traffic in the provision of internet access services and related end-users' rights" (Article 1). Zero-rating – defined as "a price of zero to the data traffic associated with a particular application or category of applications (and the data does not count towards any data cap in place)" is questionable insofar as it infringes against the activities and rights specified in Article 1.

Competition considerations are relevant insofar as "a limitation of the exercise of end-user rights is more likely to arise when an ISP [Internet Service Provider] or CAP [Content and Application Provider] has market a 'strong' market position (all else being equal) compared to a situation where the ISP or CAP has a 'weak' market position" (BEREC, 2016: para 43). The guidelines apparently presume, as argued by a number of Net Neutrality advocates (e.g. van Schewick, 2016; Lee & Wu, 2009), that increasing the range and diversity of content and applications is an objective to be pursued in its own right, regardless of the economic consequences[1]. The question of how to trade off the incentives to innovate in each of CAP and ISP activities is not addressed. As the range and diversity of content is the primary proxy for measuring "innovation in the internet ecosystem", ISP activities that materially discourage CAP entry, force their exit, or lead to ISPs "picking content winners" appear to be interpreted as per-se infringements on CAP rights as internet service end-users. The regulation is asymmetric, as there is no parallel obligation on CAPs to refrain from activities that inhibit innovation in the network elements of the internet ecosystem, or from "picking ISP winners", either in particular technologies or firm identities (Howell, 2016). Unsurprisingly, inconsistencies have emerged in the determination of acceptable behaviours in different EU jurisdictions (Elaluf-Calderwood & Layton, 2015).

1.3 Incorporating Economics Into Zero-Rating Guidelines

The BEREC guidelines therefore are insufficient as either a framework for assessing the relative economic merits of an example of zero-rating, or to assist in prioritising the cases that will justify committing resources to undertake a full inquiry. However, they provide a starting point for considering how such a framework placing total welfare as the primary consideration might be developed.

[1] Despite the economic literature on monopolistically competitive markets that indicates over-much entry is likely when entry costs are low (Carlton & Perloff, (2005).

This paper endeavours to provide insights into how such a framework might be developed. It contributes to the growing body of understanding about the economic effects of complex interactions in the Internet ecosystem. It draws upon the current literature and evidence in practice on the economics of data caps and zero-rating, and uses this to develop a framework for considering the different interactions that might be expected as the Internet ecosystem enters a more mature stage of its life cycle. This framework is intended to assist regulators and adjudicators – most specifically in the US, but also in other jurisdictions – in assessing the possible benefits and harms arising from the use of these instruments in different contexts. Its' substantive contribution is a set of questions to help guide the initial evaluation of a set of case facts, As befits a first attempt to construct an evaluation framework in a new area of inquiry, and one undertaken in a fast-moving environment where prediction is problematic, the list of questions is neither prescriptive nor presumed to be complete.. Neither does it set out to support any preconceived vision of a "neutral" or "non-neutral" Internet. Its purpose is to assist in the development of regulatory governance of an evolving Internet ecosystem enabling the delivery of the greatest possible welfare increases.

The paper proceeds as follows. Section two positions zero-rating within the current network neutrality debate. Section three summarises key insights from the sparse body of network neutrality economic literature that are applicable to examples of zero-rating. Section four the proceeds to develop a questioning framework. Drawing upon the analytical approach used by Gans (2015) and Gans & Katz (2016, 2016a), we start with a simple theoretical model for analysis and gradually introduce a range of complexities into the costs, revenues, production functions, product and consumer variations, competitive positioning, methods of competitive interaction and stakeholder strategies to identify how a specific instance of zero-rating might be expected to influence both static and dynamic stakeholder interaction and the evolution of the internet ecosystem. These understandings therefore assist in assessing the expected effects of each case upon the total welfare decision criterion. Section five summarises and concludes.

2. Positioning Zero-Rating in the Network Neutrality Debate

A small but growing academic literature exists on the subject of network neutrality. Comprehensive early reviews are provided in Frieden (2008), Schuett (2010) and Faulhaber (2011). Krämer *et al.* (2013), Bauer & Obar (2014) and Greenstein *et al.* (2016) contain reviews of more recent developments. A defining characteristic of the literature is the apparent disjunction between assertions that neutral and non-discriminatory treatment of data, applications and end-users must prevail as an inviolable Internet design principle, and the economic rationale to support such assertions. As Greenstein *et al.* (2016: 3-4) observe: "there is little support for the bold and simplistic claims of the most vociferous supporters and detractors of net neutrality. The economic consequences of such policies depend crucially on the precise policy choice and how it is implemented".

2.1 Neutrality and Non-Discrimination: Neither Synonyms Nor Substitutes

Frieden (2008) positions the origins of the network neutrality debate in concerns that network operators would manage their networks "fairly" as the Internet evolved both technologically and commercially. The form of that evolution has been conditioned by the planned and controlled path upon which the Internet – initially a heavily government-subsidised infrastructure – was gradually privatised and commercialised. As the historically highly-regulated telecommunications industry has emerged as one of the key providers of Internet access (BIAS) services (ISPs), inevitably the evolution has occurred, at least in part, in a regulatory spotlight. Hence regulatory agencies have been frequently called upon to both arbitrate on, and use their regulatory powers to ensure, telecommunications providers have acted "fairly". In addition to any regulatory obligations on telecommunications-origin ISPs, all providers of both Internet access and the applications and content provided over them have been subject to competition law obligations precluding acting in a manner that substantially reduced competition. Thus, whilst ISPs may hold a degree of market power, their ability to behave in the manner of monopolists as characterised in most economic models has been severely constrained (Howell & Layton, 2016).

Nonetheless, persistent concerns have remained that in the process of evolutionary technological and commercial innovation, holders of market power – and telecommunications-based ISPs in particular – will abuse their dominant position. In the discourse of human and political rights, this could be interpreted as acting "unfairly" by treating individuals differently based upon discernible differences – that is, discriminating. In this context, "discrimination" is not a pejorative term – it simply refers to the ability to be able to distinguish between two or more candidates. But "neutrality" and "non-discrimination" are neither synonyms nor interchangeable concepts. If discriminating between two candidates results in an "unfair" outcome, then this may be remedied by regulations or legislation requiring all individuals to be

treated "neutrally" – that is, requiring conscious ignoring of obvious differences when interacting (Bauer & Obar, 2014). Such thinking is consistent with the normative aspirations expressed in the network neutrality debate that all traffic be treated equally, regardless of its origin or destination. Schuett (2010) identifies this regulatory approach as "net neutrality as a non-discrimination rule", whilst Kramer *et al.* (2013: 796) classify it as "strict net neutrality" that "prohibits Internet service providers from speeding up, slowing down or blocking Internet traffic based on its source, ownership or destination". The bright-line prohibitions of the February 2015 Order seek to address these concerns.

However, from both economic and human rights perspectives, strictly equal treatment may lead to unequal or "unfair" outcomes. In human rights discourse, positive discrimination may be necessary, at least for a time, to remedy unfairness arising from past neutral treatment or negative discrimination (Bauer & Obar, 2014). In respect of data handling, neutral treatment of data denies the potential to derive greater total welfare that may accrue from prioritising more highly-valued data over less highly-valued data. At the very least, ISPs must engage in some traffic management to ensure their networks to operate effectively – regardless of whether or not they seek remuneration for this activity. Hence, irrespective of the distribution of the higher surplus obtained, a non-neutral approach is preferable to the neutral one (Kaldor-Hicks Criterion). For the appropriate signals to be sent to operators to undertake welfare-enhancing data prioritisation, more-highly-valued packets might be required to bring higher revenues than the less-highly-valued ones. That is, price discrimination – suppressing the neutral pricing of data transport – may be necessary for the full economic benefits of the Internet to be realised.

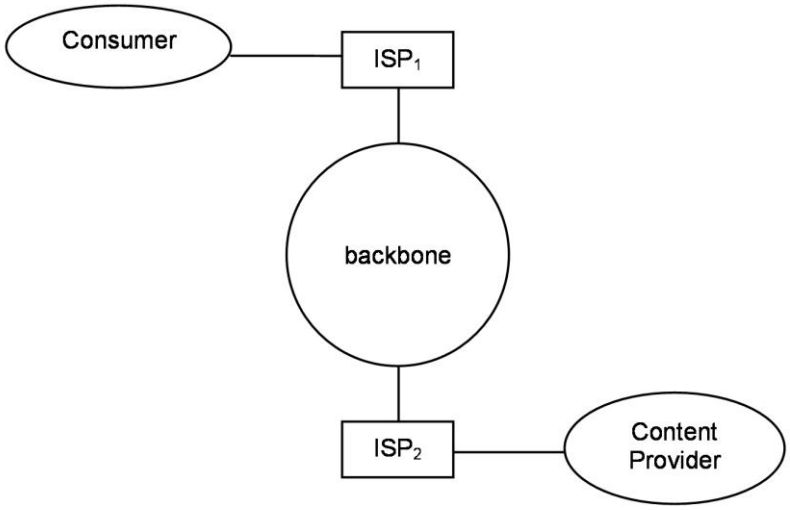
Article 3(3) of the European Regulation recognises that truly neutral traffic handling is neither possible nor desirable, so it does not prevent "providers of internet access services from implementing reasonable traffic management measures. In order to be deemed reasonable, such measures shall be transparent, non-discriminatory and proportionate, and shall not be based upon commercial considerations, but on objectively different technical quality of service requirements of specific categories of traffic. Such measures shall not monitor specific content and shall not be maintained for longer than necessary". That is, traffic management is permitted, but must not be based upon commercial considerations, even though the ensuing service quality consequences (and importantly differences in the qualities provided by different ISPs undertaking different traffic management activities) are key factors upon which end-users will likely base their decisions about which ISP to purchase Internet access services from. This suggests that for BEREC at least, "commercial considerations" encompass price-based but not quality-based competition. A similar concern appears to motivate the FCC's Open Internet Order.

Both the FCC and BEREC considerations therefore respond to a network neutrality discourse where prices and the potential for differences between them for the provision of both identical and different products and service levels to end-users have assumed a central position. However, this has not necessarily been reflected in the offers observed from both ISPs and CAPs in both jurisdictions, or in the form of the regulations adopted.

2.2 Net Neutrality vs Consumer Sovereignty: Connection Speed

Indeed, a considerable degree of price and product differentiation is observed in practice – notably in relation to the speed of the connections sold and other quality-of-service undertakings – in the relationship between ISPs and their end consumers (the upper relationship in Figure 1). Differentiating the speed of connections serving different consumers necessarily results in the traffic of the consumer with the faster connection taking priority in traffic queues over that of consumers with slower connections (Howell, 2016a). Yet neither the provision of differentiated services nor the fact that higher fees are charged for the faster connections (and consequently traffic prioritisation by user identity) have featured prominently in the net neutrality debate as non-neutral activities warranting intervention.

Figure 1: Structural Architecture of the Internet



Source: Schuett (2010), p 3.

Indeed, BEREC (2016) argues that offering a range of speeds is not per-se harmful. Rather, limiting the range of speed offers would violate the exercise of end-user rights of choice laid down in Article 1 of the Net Neutrality Regulation. Thus, speed choices are acceptable as when "applied in an application-agnostic way (applying equally to all applications), end-user rights are likely to be unaffected by these characteristics and conditions" (para 32). An alternative

statement of this observation is that it is acceptable for ISPs to charge some end-users higher prices to ensure their traffic is prioritised over that of other end-users, because it is the end-users, and not the ISPs, determining whose traffic goes to the head of the queue. Instead of addressing the preservation of neutral traffic-handling, BEREC's position is an assertion that when considering the paid prioritisation of internet traffic, consumer choice and consumer sovereignty will prevail over the interest of both ISPs and CAPs. Although ISPs decide the range of speeds offered (e.g. in their DSLAM technology choices and settings, for ADSL/VDSL networks), this echoes the primacy assigned in Wu's (2003) seminal paper on network neutrality to the rights of consumers to determine the specifications of equipment such as the dial-up modems that they wished to connect to legacy telephone networks in order to access the Internet.

2.3 Net Neutrality vs Consumer Choice and Sovereignty: Usage-Based Pricing

Given that the primacy attached to consumer sovereignty has precluded debate about customer-generated paid prioritisation, the network neutrality debate has focused largely upon the question of whether ISPs connecting end consumers to the Internet (ISP_1 in Figure 1) will be allowed to charge the providers of content sought by those consumers a fee for delivering it – that is, charge a "termination fee" characteristic of voice telephony network arrangements (Laffont & Tirole, 2001). Schuett (2010, p 2) describes the prohibition of such fees as "net neutrality as a zero-price rule". Krämer *et al.* (2013, p 797), drawing upon Hahn & Wallsten (2006) state "net neutrality usually means that broadband service providers charge consumers only once for internet access … and do not charge content providers for sending information over broadband lines to end consumers". Usage-based pricing of Internet access (notably the sale of tiered plans with increasing caps on the amount of data transferred or received) is considered by some net neutrality advocates to violate the principle that consumers pay "only once for internet access" (that is, a flat fee, regardless of the amount of use made of the connection). Zero-rating is meaningful only in the presence of usage-based pricing, because it excludes some selected (differentiated) traffic from being counted against the "cap" in the relevant consumer's Internet access plan.

BEREC (2016) includes capped data plans in the range of acceptable non-neutral activities along with speed choices so long as they are applied equally to all applications. The guidelines are silent on how consumer choices of plans with different data caps can be 'application-agnostic' when the end-user data volume demand is derived from the applications a given consumer routinely utilises (Howell, 2003). Consumers preferring applications requiring high data volumes (e.g. video streaming) will necessarily prefer plans with higher data caps than those preferring applications with lower data demands (e.g. web browsing, email).

2.4 Zero-Rating as a Neutrality Concern

Despite usage-based pricing necessarily having a stronger negative effect on the data consumption of a user preferring high-data applications than on the consumption of a user preferring low-data applications, it has been the potential that ISPs will charge either CAPs or end users differently for the delivery of specific applications or content that has attracted attention. Once again, the BEREC guidelines are instructive. BEREC defines zero-rating as "a price of zero to the data traffic associated with a particular application or category of applications (and the data does not count towards any data cap in place)". It is not necessary for the ISP to charge or remunerated by a given CAP to treat traffic associated with specific content or applications differently, as would be required to violate the Schuett and Krämer net neutrality definitions. The simple act of NOT charging the user for the delivery of specific content is sufficient to make zero-rating a concern.

Notwithstanding, BEREC appears to have created some 'safe harbors' within its provisions. ISPs are permitted to bundle provision of internet access services (IAS) with an application, so long as "the traffic associated with this application is not subject to any preferential traffic management and is not priced differently to the transmission of any other traffic", because this does not limit end-users rights (para 33). Thus, "a mobile operator may offer free access to a music streaming service application for a period of time to all new subscribers". This apparently does not constitute zero-rating as all new customers participate in the deal (although it does not address how this might discriminate against existing customers who were not offered the same opportunity). Neither is it problematic if the ISP offers zero-rating for an entire category of applications (e.g. all video or all music streaming applications) or even certain applications (e.g. its own services, one specific social media application, the most popular video or music applications), so long as an end-user is not prevented from using other similar applications from other providers and does not "lead to circumstances where end-users' choice is materially reduced in practice" (para 39).

However, "the zero price applied to the data traffic of the zero-rated music application (and the fact that the data traffic of the zero-rated application does not count towards any data cap placed on the IAS) creates an economic incentive to use that music application instead of competing ones", so is more likely to "undermine the essence of end-user rights" or lead to reduced user choice than when it is applied to an entire category of applications (*ibid*). As "price differentiation between *individual* applications within a category has an impact on competition between providers in that class", then it may be "more likely to impact the continued functioning of the internet ecosystem as an engine of innovation, and thereby undermine the goals of the Regulation than would price differentiation between classes of application" (para 45). Furthermore, "a zero-rating offer where all applications are blocked (or slowed down)

once the data cap is reached except for the zero-rated application(s)" would count as an infringement not because of its potential effects on end-user rights or incentives to use a specific application over another, but because it would constitute a breach of Article 3(3) requiring all traffic to be treated equally without discrimination, restriction or interference and irrespective of the sender or receiver, the content accessed or distributed, the applications or services used or provided, or the terminal equipment used (para 52).

It is therefore quite unclear how, within BEREC's guidelines, the various objectives can be traded off against each other where zero-rating is concerned. It is possible to take from reading these guidelines that discriminate traffic management is not permitted, except where it is necessary for the efficient management of the network, but not if it means treating traffic associated with one application or class of applications differently from another, except in the case where failing to discriminate leads to reduction in end-user choice, where not only is it possible but in fact necessary to create the circumstances where one users' traffic (which may or may not be of a particular application class, but is most likely to be a function of a particular application choice set) will be prioritised over that of another user. As Katz (2016) and others have observed of the FCC's Open Internet Order, they are confusing, contradictory and lack underlying economic logic. Despite claiming to be guided by principles of promoting consumer choice and sovereignty, they are likely to lead to a number of unintended consequences that will harm, rather than assist, that objective.

2.5 'Saving the Internet' via 'Net Neutrality' is Neither 'Neutral' Nor 'Desirable'

It is noted that considerable advocacy surrounding the introduction of the February 2015 Open Internet Order was based upon strong calls to "save the Internet" from changes to ISP charging practices (e.g. Save the Internet, 2016). These calls presumed a normative superiority of arrangements with flat-rate tariffs and no interconnection fees, and asserted that they comprised a crucial "design element" of the original Internet. Changing ISP charging practices were portrayed as an assault on the Internet itself. Regulation was invoked not on economic grounds but using arguments reminiscent of defence of a threatened territory or preservation and conservation of a species of flora or fauna threatened with extinction. However, Howell & Layton (2016), drawing on (amongst others) Frieden (2008) and Leiner, *et al.* (2003), propose that these ISP charging strategies were not explicit design features of the originally-privatised Internet, but arose as pragmatic and efficient responses to a range of commercial, technological and usage characteristics prevailing at that time. Furthermore, those engaged in the early commercialisation of the Internet anticipated changes in pricing structures would emerge as the Internet evolved in its new environment.

If the intention of some neutrality advocacy is for regulations to be invoked to prevent the internet ecosystem from evolving in a particular direction, then it would appear necessary

to justify the economic benefits of such intervention against a counterfactual of not intervening It is not sufficient to justify intervening simply to prevent any changes from occurring. In an evolutionary context, preventing change removes dynamism and imposes stasis – i.e. death – on a system. In an economic context, suppressing dynamic evolution – albeit even temporarily – serves to cement in place a particular set of distributions of the gains from activity achieved from the arrangements observed in the static frame. This is in direct contravention of the objective to preserve the Internet ecosystem, as opposed to one particular subset of it, as an engine of innovation.

3. Positioning Zero-Rating in the Economic Literature

Despite the apparent concerns expressed recently in the net neutrality debate about the potential economic harms of zero-rated offers (for example, since T-Mobile began offering zero-rating for streamed content from its Binge-On platform, and Comcast did likewise with Stream TV), there is comparatively little analysis of them in the economics literature on network neutrality.

As identified above, the predominant concern to date has been the potential for ISPs to charge "termination fees" to content providers in order to deliver their content to specific end-users, either with or without concomitant offers to tie termination fees with giving some traffic priority, or in the absence of such fees, blocking (refusing to deliver) or throttling (slowing) selected traffic relative to other comparable traffic for which fees have been paid (e.g. Choi & Kim, 2010; Economides & Hermalin, 2012; 2015; Choi, Jeon & Kim, 2014; Kourandi, *et al*., 2015; Peitz & Schuett, 2015; Gans, 2015; Gans & Katz, 2016; 2015a). Three possible motivations have been suggested for ISPs to charge content providers: to facilitate in collusion with selected affiliated providers to prioritise the delivery of their content over that of their rivals, to foreclose content market competition; the exertion of market power in order to appropriate higher surpluses from content providers (notably when those content providers have some market power and charge end-users for content accessed); and to recoup some of the costs of investing in the substantial capacity upgrades necessary to deliver the substantially-increased traffic volumes associated with a handful of applications – notably video streaming services such as Netflix (Greenstein, *et al.*, 2016).

The small but growing number of papers addressing these concerns have contributed to a richer understanding of the issues involved. However, as Greenstein *et al*. (2016) observe, the findings of each paper are highly contingent upon the specific contexts that they model. Neither are the assumptions made in them necessarily reflective of the actual interactions taking place in the actual markets in which the regulations apply. Therefore it is not simple to translate their findings into coherent and meaningful policies. For example, models that assume ISPs hold and can exert market power over end consumers are not necessarily helpful in

understanding what might occur in contexts where access regulation renders fewer barriers to new ISPs entering to compete for surpluses generated by the charging practices of their established rivals (e.g. fixed-line infrastructure markets in Europe, Australia and New Zealand). Likewise, the options created by access regulation and network neutrality provisions may alter both the incentives to form alliances and mergers between ISPs and CAPs, and the potential competitive harms that may arise from such activities, but not necessarily in a consistent manner.

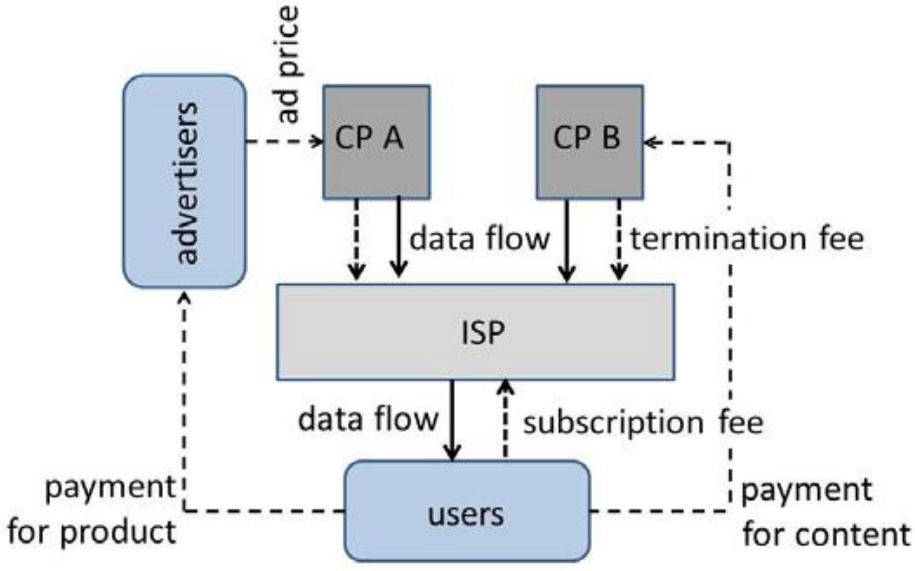
Furthermore, the majority of the papers considering selective payment for end-user application and usage model the payment as a single fixed fee paid by the CAP (e.g. Choi & Kim, 2010; Guo, *et al.*, 2013; Economides & Hermalin, 2014). These papers focus on the levels of congestion faced by ISPs as data volumes increase, and the potential for ISPs to maximise their profits by charging selective CAPs to prioritise their traffic in order to maintain application quality delivered to end-users. They also assume that CAPs do not charge end-consumers for their content. The CAP payments lead to differential service quality delivered to end-users of those applications, but do not alter the prices paid by them. Furthermore, as they concentrate on the strategic choices of ISPs and ISP investment incentives, these papers do not address the concomitant effects that application-based differentiated service quality may have on either competition in the CAP market, or the incentives for CAPs to invest new content and application development. Conversely, papers focusing on the competitive effects in CAP markets from subsidising end-users' data consumption (e.g. Ma, 2014) do not address the concomitant effects in ISP market competition and investment. Their focus is on the effects in the markets for advertising revenues. Cho, *et al.* (2016) appears to be the first (and to our knowledge so far the only) paper to explicitly model differential price effects for end-users of different applications, and takes account of heterogeneity in consumer application preferences when selecting the applications that generate data traffic. However, as with most other papers in this category, the authors presume that CAPs subsidising their end-users' data consumption receive no other revenues from them, so must rely only upon advertising revenues to generate the subsidy funds.

Each of these papers provides some interesting insights into some specific interactions that occur within the Internet ecosystem because, for the most part they have been developed to address one specific set of behaviours that may or may not be a potential breach of one particular tenet of the net neutrality debate. Consequently, they are not especially helpful for generating a set of principles or guidelines to assist regulators and adjudicators in identifying the relevant economic trade-offs across the wider internet ecosystem that they are required to assess when evaluating a specific case of zero-rating brought to their attention. In this regard, a new approach to considering the issues pioneered by Gans (2015), Gans & Katz (2016; 2016a)

and Greenstein (2016) promises to be more helpful. These papers start with a simplified structural model of the internet ecosystem which identifies the key stakeholders, the interactions between them and the cash flows generated. The structural model is systemic, as it identifies the potential for stakeholders to interact with each other in different capacities at different times, thereby capturing some of the interconnectedness that the situation-specific economic models overlook. The primary intention of these models is not to identify an optimal set of arrangements, but rather to identify how spillovers and externalities travel through the ecosystem. The authors start with a very simple model that connects all the relevant participants and transactions in the system, and then proceed to add complexities to subsets of the transactions. The effects of these changes can then be traced via the interconnections of the simple model to the parts of the ecosystem that are not explicitly considered in the situation-specific economic models that have to date characterised the net neutrality literature. The simplicity of the original model readily enables identification of the broad range of trade-offs that will have to be considered by regulators and adjudicators when considering the specific facts of the case they are required to assess. Situation-specific economic models can then be created to identify and evaluate how these trade-offs will play out.

The use of a simple model to identify these tradeoffs is illustrated using Figure 2.

Figure 2: The Current Internet Ecosystem



Source: Greenstein *et al*. (2016: 12).

In this conception, Content Providers (CPs) are generally held to be competing with each other to attract users to view their content. CPs derive their revenue from a mix of advertising revenues and content payments from IUs. The more IUs the CP can attract, and the higher the number of views of CP content each IU makes, the higher is the expected advertising revenue for that CP. If advertising revenue is high enough, the CP need not charge IUs separately for the content viewed. Thus CPs also act as two-sided platforms linking IUs and Advertisers. However, assuming there is a limit to the total amount of advertising revenue available, the more intense is the competition between CPs, the more thinly will advertising revenues be spread across CPs, and the more likely it is that the CP will have to charge IUs to access their content. As there are limits to both the number of potential IUs and the number of visits they can make to content sites of any CP, raising charges to users will thus lower both the expected number of IU customers and the expected number of visits each IU makes to any given CP. Expected advertising revenue thus decreases (Gans, 2015; Kourandi, et al., 2015). The greater is the share of a CP's revenue derived from advertising, the greater is the threat to its business from similar advertising-focused CPs. Consequently, CPs have a keen interest in both absolute and relative tariffs charged by ISPs. It will not always be the case that intense competition in ISP markets will be in CPs' interests.

ISPs provide the technical means by which data travels from CPs to IUs. They typically charge IUs a subscription fee, which may be a flat fee, a variable fee per packet transferred or a mixture of the two. The greater are the number of IUs and the higher is their usage, then the higher will be expected ISP revenues. However, ISPs hold 'terminating market power' over their IUs (Economides & Täg, 2012; Krämer *et al.*, 2013) as – assuming each IU has a commercial relationship with only one ISP – they control the ability for CPs' data to reach any given IU. Hence ISPs can potentially charge CPs a delivery or 'termination' fee to deliver the data. A monopoly ISP is able to maximise its profits by using termination fee revenues to offset subscription fees to obtain the optimal number of IUs and page views. Gans (2015) and Greenstein et al (2016), however, show that although the ISP may levy termination fees, the power to do so derives from the agency to do so conferred upon it by the IU when opting to use a given ISP to access the Internet. Absent competition, the ISP can balance subscription and termination fees to maximise profits. But with competition, termination fees will be passed on to IUs in lower subscription fees in order for the ISP to retain and attract IUs. If ISPs and IUs can extract all profits in this manner, then one of the key assumptions underpinning net neutrality advocacy – that content providers rely upon implicit subsidies from ISP revenues (i.e. by not having to pay a termination fee, as might otherwise be expected in optimal two-sided pricing) to incentivise new content creation – will not hold.

Gans (2015) and Gans & Katz (2016, 2016a) then proceed to illustrate how successively changing the starting assumptions of the simple model adds complexity without losing sight of the key trade-offs embedded in the systemic interactions across the multiple two-sided platforms. In the initial model, consumers are homogeneous, the content of the two providers are substitutes and ISPs are monopolists. In this basic setting, they show that when CPs charge IUs, it makes no difference to total welfare when ISPs are allowed to charge CPs – all that is affected is the distribution of the gains between CPs and ISPs. Likewise, introducing competition between ISPs does not alter the total surplus, but leads to the ISPs' share being appropriated by IUs. Introducing a weak form of net neutrality (preventing price-based discrimination in the fees charged by ISPs to either CPs or IUs) has no effect on the equilibrium outcome, and even strong neutrality (preventing price discrimination in both CP and IU relationships) does not always affect the equilibrium outcome.

The effect of strong neutrality rules depends upon the degree to which content offerings are perfect substitutes or are differentiated, and the degree to which end users' content preferences are differentiated. Where a single representative consumer prefers one content over another (i.e. they are not substitutes, strong neutrality leads to every consumer choosing the optimal provider, so the non-optimal one must exit the market. Weak neutrality on the other hand allows the ISP to set a pricing strategy that allows both provides to remain in the market – that is, it need not use prices charged to CPs to exclude either provider. Content provider exit occurs because of consumer preferences and not ISP pricing strategies. The extent to which the ISP and the remaining content provider share the surpluses will depend upon whether the price charged by the ISP to the CP is allowed to fall below zero (i.e. can the CP induce the ISP to pay it for the right to carry its content). Setting termination fees at zero, but allowing the ISP to vary prices to IUs can lead to inefficiently too much content investment. Intuitively, this occurs because there is no link between the marginal benefit of increased content quality to consumers and the marginal benefit to the ISP, which forces the consumers' content choice. Allowing consumer heterogeneity alters the equilibrium, but strong neutrality rules will harm welfare under some circumstances, relative to the case of no neutrality rules or weak neutrality, again because the neutrality rules prevent the conversion of relevant information into price signals passing between ecosystem participants.

Whilst the use of these simple models does not alter the complexity of the trade-offs that are required to understand the full effects of different neutrality rules in specific circumstances, they do raise a number of different considerations that have not featured so far in the net neutrality debate, so are not explicitly considered in the BEREC guidelines. Gans (2015) and Gans & Katz (2016, 2016a) highlight the extent to which both content and consumers are differentiated can have a powerful influence over the effects of different net

neutrality rules on total welfare. For example, the primacy given to consumer choice in the range of ISP prices and services offered in the BEREC guidelines may have negligible effect on incentives for efficient (welfare-enhancing) content creation if the predominant source of content subsidy is from advertising revenues and not the absence of a termination fee in providers' contracts with ISPs.

4. Analysing Zero-Rating Using Simple Economic Models

In this section, we follow the analytical approach used by Greenstein, Gans and Katz to explore how incremental extensions to simple economic models allows insights to be gained into the likely effects of zero-rating on economic welfare. It is not our intention with this analysis to evaluate any specific example of zero rating. However, we will draw upon examples to illustrate where it may be necessary for regulators or adjudicators to consider complex trade-offs when assessing some specific cases. In this way, we will demonstrate how the complexities of the internet ecosystem render the BEREC guidelines both overly simplistic and potentially misleading in assessing the effects of zero-rating. We propose instead that our questioning framework facilitates both a clearer understanding of the interactions occurring within the internet ecosystem, and the more appropriate selection of the relevant economic models for assessing the effects. This is consistent with, but does not restrict the case-by-case analytical approach adopted by the FCC, provided maximising total long-run welfare is the primary objective. However, it does challenge the foundations of the BEREC approach, which are predicated upon promoting competition as the primary objective.

4.1 Perfect or Imperfect Competition?

Our starting point is the primary assertion of net neutrality advocates that zero-rating allows ISPs to 'pick winners' in application and content (henceforth 'applications') markets by using subsidised access prices for selected applications delivered to specific end-users, with the intention of foreclose unsubsidised rival offerings. Reducing the number of applications is considered by net neutrality proponents to be antithetic to the objective of promoting the internet ecosystem as an engine of innovation. This argument, proposed initially by Lemley & Lessig (2001), presumed innovation at the 'edge' of the ecosystem unconditionally dominates innovation at the core. It has been used recently by van Schewick (2016) to question the efficacy of T-Mobile's zero-rating of content on Binge-On. It has also been raised by critics of Facebook's Free Basics.

To gain insights into the extent to which zero-rating of selected applications may lead to the foreclosure of others, we will begin with a simple, one-sided model of perfect competition, and then extend the analysis across both internet access and application platforms in the internet ecosystem. Under perfect competition, no market participant has any market power. The price at which a given homogeneous product is exchanged is set by the intersection of supply and demand. The former reflects the costs of supplying the product, and the latter consumers' willingness to pay. There are no transaction costs, barriers to entry or exit and perfect divisibility of output. The price at which the marginal unit is exchanged is determined by the equality of the cost of producing that unit and the marginal consumer's willingness to pay for it. All participants also have perfect information about each other's costs and prices.

This combination of circumstances almost never occurs in real markets. However, the ways in which real markets vary from those hypothesised in models of perfect competition are instructive for considering the primary claim of net neutrality advocates that all instances of zero-rating of ISPs' charges to their consumers are likely to be harmful to the interests of both applications providers and consumers, because they pose barriers to innovation in the application zone of the internet ecosystem.

4.2 Demand for Internet Access is a Derived Demand

Before proceeding further, it is important to understand that end-users' demands for ISP access services are derived from their demands for applications. An ISP connection is of no value to a consumer if it is not used to access internet applications. The value of the connection is therefore dependent upon the value the consumer places on the applications accessed. The ISP may supply some of these applications, but for the most part, consumers' value of the connection is contingent upon being able to access applications not provided by their ISP. This nontrivial observation illustrates why, for the most part, ISPs do not have strong incentives to impede their consumers' access to the huge range of applications available on the Internet.

If ISPs do not supply applications themselves, then all else held equal, restricting or impeding access to specific applications reduces the value consumers derive from the connection, so will reduce the number of consumers who will buy a connection at any given (fixed) price, and the utilisation of those applications, which generates revenue under metered billing arrangements. Furthermore, fewer connections and lower usage reduces incentives for application creation, lowering both the revenues application providers anticipate from both subscriptions and other sources (e.g. advertising) and the expected future benefits that consumers anticipate from new and unanticipated capabilities provided by applications yet to be developed. This interdependence

4.3 Internet Users Have Extremely Heterogeneous Demands

ISPs can charge consumers a flat fee, a usage-based fee or a combination of the two for internet access. Consumers' internet access purchases are determined by trading off the fixed price paid for access and any usage charge against the benefits of accessing and utilising applications. Menus of two-part tariffs bundling access and usage charges are generally welfare-enhancing relative to a single flat-rate or solely usage-based tariff as they allow users with different valuations associated with different usage levels of even a single application to self-select the tariff that gives them most surplus. This is reflected in BEREC's guidelines in the primacy given to consumers' choices. A zero-rated tariff applied to a specific application is simply a tariff with no usage-based component – that is, a flat fee. Flat fees are most advantageous for those with the highest expected usage, as they will utilise it up to the point where no further benefits will be obtained. This is necessarily more than if usage is charged at marginal cost

(noting that network congestion is a significant externality proportional to utilisation that is imposed by users when utilising applications). If the higher costs associated with higher usage levels are to be recovered in user fees, a single flat-rate tariff will be higher where usage is higher that when it is lower. Metered tariffs (including plans with flat-rates within a given data cap, that rise as the data cap increases), are an efficient means by which ISPs may recover revenues from each consumer rising in proportion with the costs that usage imposes on the ISP (including the costs of congestion that lower service quality for all users).

However, metered tariffs will arise in practice only if consumers are heterogeneous in their valuations of Internet application usage. If all consumers value their connections identically, then there will be one tariff that is efficient for all users, and there will be no incentive for ISPs to offer any other tariff. Consumers' valuations of internet application usage is inherently heterogeneous because different users will prefer to use different applications for different purposes. Some will prefer applications requiring high usage (e.g. video streaming) whilst others will prefer applications with lower resource demands (e.g. web browsing, email). Even consumers preferring a single application will vary in their use of it due to personal resource constraints – for example, time to watch streamed video and the cash to pay for the connection.

Consequently, internet access as sold by ISPs is not a homogeneous good – it varies with the application preferences of the consumers using it. Consumers with higher valuations for a single application will consume more resources than those with lower valuations. If metered tariffs are intended to recover higher revenues from higher-using consumers of a single application, then offering a zero-rated tariff for that application is inconsistent with the ISP's objective to recover its costs in usage fees. Assuming that the ISP does not recover the revenues lost from zero-rating application usage from the application provider, and it costs the same to deliver a unit of each application to the end user, then it is strategically illogical for the ISP to charge for the usage of one application and zero-rate usage of the other. Costs remain unchanged, but revenues will fall. Hence, in the simplest case, as zero-rating by an ISP discounts revenues received from selected end-users on the consumer side of the ISP platform, it must necessarily be associated with compensatory revenue streams – for example, higher fees charged to non-selected users, charges to application providers on the CAP side of the platform, or revenues from other sources, such as taxation or advertising - if in the long run the ISP wishes to remain solvent.

4.4 Zero-Rated Tariffs Alone Won't Foreclose Differentiated Applications

In net neutrality discourse, the logical inconsistency of an ISP zero-rating some applications could be explained if the ISP has a strategic reason to steer its consumers away from using one application and towards using another – for example, if it is paid by an application provider to

do so, in order to foreclose the non-paying application from the market. It might also be plausible if the ISP is also the application provider, and uses its market power over its consumers to incentivise them to use the aligned application over the non-aligned one. The primary reason to do this, according to the net neutrality discourse, is to foreclose competitive rivals. However, the foreclosure strategy will be successful only if the two applications are close substitutes. As shown by Gans & Katz (2016a), if the two applications offer materially different value propositions to end consumers, then the zero-priced application will not be able to force the positive-priced one from the market so long as there are consumers who prefer the positive-priced one over the zero-rated one by more than the discount embedded in the zero-rating offer. This is the same situation faced by broadcast television and newspaper providers. Free-to-air television and free newspapers have not foreclosed pay television and newspaper subscriptions. Indeed, some consumers willingly consume both, even when some of the content overlaps, because the additional value offered by the pay version is sufficiently high enough that it overcomes the price differential. Arguably, the presence of the two different newspaper forms has led to greater content variety, with subscriber newspapers providing a professional journalist-based news service, and free newspapers relying more upon content generated by readers (e.g. local school and sports reports) and advertisers.

This leads to our first question to be posed by regulators and adjudicators evaluating zero-rating offers.

Question 1. What very close or perfectly substitute applications accessible over the ISP's connection, costing the same to deliver, are likely to be foreclosed by the zero-rated application(s)?

The implication of this question is that the closer are the non-zero-rated application(s) to the zero-rated one(s) in the perception of the end-users, then the more likely it is that the non-zero-rated applications will be crowded out. However, there are very few applications meeting this requirement that are truly close substitutes. For the most part, content distribution networks such as Netflix and Hulu are not close or perfect substitutes for each other because they contain different bundles of content for which end users have distinct preferences. The applications themselves are differentiated, even it costs the ISP the same to deliver a Hulu movie and a Netflix one of equivalent specifications. If a consumer preferring Netflix is not prevented from paying the higher usage fee to watch Hulu content if the content available only on Hulu is sufficiently highly-valued, then Hulu will not be foreclosed, even in respect of the subset of Hulu-preferring consumers on the discriminating ISP's network.

It might be a concern, however, if the applications in consideration were, for example, two identical cloud storage applications. The zero-rated application will have an unequivocal advantage over the non-zero-rated one, leading to all consumers with a non-zero valuation of

using cloud storage opting for the lower-cost one. However, for foreclosure to occur, it is necessary for the applications to be undifferentiated – that is, homogeneous products. Foreclosure of differentiated products will be a function of the degree of differentiation – the more similar they are, the more likely it is that foreclosure will occur.

4.5 Equalising Prices Conceals Underlying Cost and Valuation Differences

The logic applied in this simple illustration leads to the same conclusion as Gans & Katz (2016a) that without some non-neutral pricing signals, over-much (inefficient) investment in application variety is possible if equalising the prices faced by consumers and application providers conceals underlying real differences in costs and user preferences.

Assume now that the two applications are perfectly homogeneous, but one actually costs less to deliver than the other. This could be because the ISP has been able to customise the delivery of one application within its own networks so that it costs less (or causes less congestion) than an otherwise equivalent one that has not been customised. It could also be that one class of applications can be processed via a different operational process that is less costly. The latter was the case in Australia and New Zealand in the mid-1990s, then the internet was first becoming popular. At the time, international bandwidth was limited due to technical constraints on the PACNET sub-oceanic cable, and due to asymmetric data flows, Australian and New Zealand ISPs purchased PACNET capacity under transit arrangements rather than peering. Traffic to and from end consumers requiring PACNET capacity was more costly to handle than traffic handled under local peering arrangements. The original retail internet plans metered international traffic by volume, but offered unmetered (i.e. zero-rated) local traffic.

In this instance, zero-rating low-cost local traffic but metering high-cost international traffic reflects real differences in underlying costs. Zero-rating that diverts consumers' usage of substitutable applications towards lower-cost substitute applications raises efficiency. However, discounting local applications did not crowd out content from foreign origins – indeed, foreign content and applications were overwhelmingly preferred by end-users, even it was metered. The free option was subsequently withdrawn in New Zealand, because it accounted for around only 5% of internet traffic.

This leads to our second question for regulators and adjudicators:

Question 2. Does usage of the zero-rated applications actually cost the ISP less than equivalent usage of non-zero-rated applications?

If the answer to this question is 'yes', then zero-rating would not lead to a perfectly efficient outcome, but would likely be less harmful to total welfare than the alternative of requiring all usage to be charged at a single price. Under the two-price arrangement, more usage than efficient would be made of the low-cost application, and the high-price usage tariff would have to be above cost to subsidise the additional low-cost usage. Arguably, this could lead to some

low-cost applications surviving that would not otherwise be viable if their usage was charged at cost – that is, Gans & Katz's inefficient over-supply of application variety.

However, the alternative of a single positive usage price that does not signal the different underlying costs will lead to more usage of the high-cost application than if it was charged at cost, which would have to be subsidised by users of the low-cost application. Increasing the price of using the low-cost application above its cost to subsidise the high-cost usage leads to less usage of the lower-cost application, and at the margin some consumers will give up their internet connections entirely because they no longer receive utility higher than the combined price of access and usage. Without the fixed revenues of these low-cost consumers to offset the higher usage costs of the consumers paying below cost, the average usage cost per unit of traffic handled increases, leading to even higher usage fees and a second depressing effect on the usage of and fees generated by low-cost users. That is, a 'waterbed effect' emerges (Katz, 2016).

Hence, zero-rating of applications with lower costs than non-zero-rated applications is not equivalent in its effects to zero-rating applications with the same costs as their zero-rated counterparts. The difference is material. In a perfectly competitive market, it is necessary for the price signals associated with lower costs to be sent to consumers so that efficiency-raising changes in purchasing behaviours can take place. Concealing information about cost differences (e.g. by averaging the prices for two or more applications) prevents consumers making efficiency-raising choices.

4.6 Differentiated Price and Product Offers to Low-Valuers

We now turn to another way in which 'averaging' the prices paid for different applications can prevent signals about consumers' preferences being conveyed and acted upon. Pro-net neutrality advocates frequently argue that it is 'unfair' to provide free or discounted access to a narrow range of internet applications or applications with some functionality removed, but to charge a higher fee for unrestricted access to the unrestricted applications. Their logic is that consumers of the restricted offers cannot participate equally with unrestricted consumers in accessing the full potential benefits of all the applications and content available on the entire internet. The ability the have unrestricted access to the entire internet content is deemed a 'right', so any arrangement that allows differentiated access to that content is seen as an infringement of that right. Zero-rating that waives all access charges in exchange for limiting the range and functionality of applications offered is therefore an infringement against the presumption that the only 'fair' arrangement is for all internet users to face the same access terms. This variant of zero-rating does not depend upon internet access and usage being charged separately. Free Basics, where potential internet users in developing countries are offered free

access to a restricted range of applications, but can access the full versions when paying a monthly internet access subscription, is frequently cited as such an infringement.

Because a user's internet activity is a bundle of access and application usage, zero-rating access to a restricted application set is in principle no different to an application provider choosing to make some content available freely, and releasing other content only when some other obligations have been met. This could include paying a cash fee, but equally could include sharing specific user information (e.g. an email address or personal information) that can subsequently be monetised. On face value, as it is the access provider, and not the application provider, who sets the tariff and faces reduced revenues as a consequence, from zero-rating access to the restricted applications, it appears not to be a rational choice unless the access provider is either also the application provider, or compensated by the allocation provider for the revenues foregone (the case of offering a 'free sample' to reduce information uncertainties is covered separately). That is, although the revenue consequences appear to be faced by the ISP, it is the application provider who makes the decision about restricting the application range to self-selecting end users. Preventing application providers from offering these discount arrangements appears at odds to the net neutrality argument that edge providers and not ISPs exercise control over internet content. If the range of content is restricted by applications providers – for example to foreclose other application providers – then that is a matter for antitrust and not internet access regulation to address.

Furthermore, the presumption that all end users should pay identical prices to access the same applications ignores economic realities. The expectation that all consumers pay the same price for a product is an artefact of perfectly competitive markets. If all consumers pay the same price, then those with higher valuations of the bundle receive more surplus than those with lower valuations. Perfect equity in access prices for a homogeneous good cements in place extreme inequities in surplus distribution. Price discrimination (same price for the homogeneous good) effectively transfers surplus from high-valuers to low-valuers and leads to higher total consumer numbers without reducing total welfare. Furthermore, where scale economies are present, then total welfare increases as well. Product differentiation (e.g. offering a subset of functionality for a lower price) leads to higher consumer numbers in total than with a single price for the undifferentiated good. Price discrimination and product differentiation therefore both appear consistent with (or at least not per se harmful to) increased product variety, larger total numbers of internet users and ongoing innovation in the internet ecosystem. That does not mean that the practices might not, in some circumstances, lead to negative outcomes. Rather, it reinforces the merits of a case-by-case analysis.

Price and product differentiation are important ways of enabling individuals with low valuations of internet use, or facing significant financial constraints, to become internet users,

The former case occurs in mature markets, when the last-remaining individuals have not yet connected because the value they place on the connection is less than even a very modest single price charged. The latter case arises in developing economies, where income constraints pose significant barriers to purchase for large numbers of individuals. Whilst subsidising connection fees through a tax and redistribution system may induce purchase in the former group, subsidising via applications may be more effective because the application is the primary determinant of the value derived. It also offers a superior means of subsidising in the latter case, because surpluses generated by users in developed economies can be transferred via the application and access bundle to subsidise those in developing economies. Thus, wealth transfers across national borders can occur without the need for government intervention.

This gives rise to our third question for regulators and adjudicators:

Question 3. Is zero-rated access to a subset of applications primarily intended to increase the number of individuals using the internet?

The purpose of this question is to separate out instances of zero-rating that are more likely to lead to positive network effects arising from larger total numbers of internet connections from instances that may arise from other motivations – for example to change the range and usage of applications by individuals already purchasing internet connections.

4.7 Who is Questioning the Efficacy of Zero-Rating?

The presumption that all internet users should pay the same price for access to the same applications prevents the surfacing, utilisation and sharing of private information that could lead to more efficient outcomes. Requiring only one price for one standard access product requires transactions to proceed as if all consumers are homogeneous, when in fact their preferences are inherently heterogeneous. Artificial imposition of the homogeneity assumption via regulatory rulings is a substantial intervention infringing upon the rights of consumers to use their information as they see fit – and could in some circumstances be interpreted as a regulatory taking, in the same manner as restricting the range of products an ISP can offer or requiring it to make regulated products available at regulate prices is also a regulatory taking In principle, such takings should proceed only when there is a clear empirical case that the net harm that would result if the right is not taken is both substantial and difficult to compensate (Wilkinson, 2008). It seems somewhat paradoxical to claim to be preserving consumer sovereignty by taking real rights from real consumers to uphold an artificial 'right' for consumers to participate in to an outcome that exists only in a theoretical model of perfect competition.

Furthermore, imposing the assumption of consumer homogeneity reduces the amount of information available to both ISPs and CAPs to customise their offerings to individual consumer preferences. Information that would otherwise have been efficiently signalled or

screened in customised offers can only be obtained subsequently by other means – inevitably with higher transaction costs. In the long run, this would seem to impose impediments to, rather than incentives for, the development of new applications and contracting arrangements. That is, banning zero-rating because the practice may pose entry barriers for new application providers must be balanced against the entry barriers that will be created if information about underlying consumer heterogeneity that would be efficiently signalled, screened and shared if zero rating proceeds cannot emerge due to regulatory intervention banning the practice.

Whilst banning zero rating has been justified by the potential for ISPs to raise the costs for new application providers, it is equally plausible that banning prevents both allocation developers and ISPs from both learning about and creating offers that cater to these underlying differences. Thus, existing ISPs and application provides might prefer the information not to be surfaced if in doing so, opportunities were created for new entrants to take advantage of consumer heterogeneity to create new offers, attract consumers away from the exiting providers and appropriate a disproportionate share of the new consumers yet to purchase internet connections. Likewise, existing internet consumers obtaining high surpluses under a single price might be unwilling to share those surpluses with new consumers who will participate only with implicit subsidies.

This gives rise to our fourth question for regulators and adjudicators:

Question 4. Who has requested that an instance of zero-rating be investigated?

If the request has come from existing ISPs, then it is plausible that the motivation may to foreclose competitive entry by rival ISPs. If it has come from existing applications providers, then the motivation may be to foreclose competitive entry by new applications provides. If it comes from existing internet consumers, then the motivation may be to lock in existing surpluses and not have to share them with new or future internet consumers. On the other hand, if the request to investigate has come from potential ISPs or application providers then the claim that it creates an entry barrier may be credible. It seems most unlikely that a non-consumer would ask for an inquiry about the legality of a zero-rating offer that would cost less than the alternative price. Similarly, it is also unlikely that a low-valuing existing consumer who would be better-off using the restricted zero-price offer would request an inquiry.

4.8 Zero-Rating Can Reduce User Search Costs and Address Incumbent Advantages

So far, we have used the model of perfect competition as the primary model for assessing interaction in both the ISP and CAP parts of the internet ecosystem. However, in using this model, we have found it to have some limitations, which have formed the basis for our four preceding questions. Specifically, it presumes homogeneous products, perfect information, no transaction costs, no externalities, price-taking participants, and perfectly divisibility of output

– none of which occur in the real markets. It also focuses on only one time period, so does not lend itself well to the dynamic interaction that typically occurs in the internet ecosystem.

In markets with heterogeneous products, consumers with different preferences, and information asymmetries that make it costly, if not impossible for consumers to identify the attributes of the products or the fit with their preferences before they have been consumed, a more appropriate model for analysing interaction is monopolistic competition. In this model, within a range of products (ISPs or applications) there will be one that will be the best match for a given consumer with given preferences. At any given price, this product gives the consumer the highest possible surplus.

However, the consumer cannot identify in advance which is the best match. Nor can the provider accurately identify the best consumers for the offer. The consumer can select one offer at random - so long as the surplus from this purchase is not negative, the consumer has gained at least some increase in surplus. Where the consumer will use a service multiple times (or make multiple purchases), the gain from purchasing the same product/service is known. There may be a better match available (higher gain) from buying a different product next time – but there is also a risk that the different product is a worse match than the existing one. The consumer could have had higher surplus if instead the first product had been purchased. There may also be switching, learning and adjustment and other investment costs associated with each product. Buying from a second supplier means a second investment in these costs – which is avoided if second and subsequent purchases are made from the first supplier. Together, these comprise 'search costs'. The larger are the search costs, and the smaller is the expected benefit of the second product over the first, the less likely it is that the consumer will try to find a better match, even though there is definitely a better one out there. Thus, high search costs lead to suppliers having some market power over their existing customers – akin to monopoly – even though there are many different variants of the product -competitors – available for consumers to choose from.

Almost certainly, the markets for internet application adoption and usage are monopolistically competitive. Customers make investments in using specific applications (learning costs, emotional investments etc) that make them reluctant to try new variants. When a new application enters a market where customer preferences are already well-established, overcoming these high search costs is likely one of the most significant barriers to be faced.. The more mature is the application market, the more established are these preferences and the harder it will be to overcome them. Even if the new product is superior to all others in the market, customers will be reluctant to try it, because they do not know that it is better for them until they have tried it. If the same price is charged for the new and existing products, the new product will not attract any new customers, because of the high search costs customers face. In

this case, the only way that the new product will attract new customers is by charging less than the existing products – that is, undertaking to meet the search costs incurred by the customers. For this reason, new products in this market are typically introduced with free trials.

However, if a new internet application is offered free of charge to consumers, because the costs are recovered from advertising or other sponsored revenues (e.g. donations, tax funding), it is not possible to discount the application cost to encourage switching. The only way that potential customers' search and switching costs can be reduced is by reducing the internet access charge. Hence, zero-rating may be the only viable way of inducing existing consumers to try a new product. Not being able to offer zero-rating thus constitutes an entry barrier to new applications seeking to compete with established ones. Just as in question four, it will be existing applications providers, and not new entrants, who would prefer that zero rating not be allowed. However, it is important to note that there are two different reasons for coming to this conclusion.

This gives rise to our fifth question for regulators and adjudicators:

Question 5. Do consumers of the zero-rated application and its rivals make payments to applications providers separate from their payments to ISPs?

If the answer to this question is 'yes', then the appropriate models to use to assess the economic consequences are those of the form provided in Gans & Katz (2016a). However, if the answer is 'no', then it will be more appropriate to use models of the form developed in Cho, *et al.* (2016). However, it is noted that neither of these models take account of the effects of unmetered usage on the costs of congestion faced by ISPs (as per Choi & Kim (2010) and Choi, *et al.* (2014)). This suggests that more theoretical work is needed to explore the full range of trade-offs arising in each of these two cases.

5. Summary and Conclusions

In the introduction to this paper, we identified that requirement of the FCC's February 2015 Open Internet Order that instances of zero-rating be assessed on a case-by-case basis was a pragmatic recognition of the fact that it is extremely difficult to identify with any certainty the circumstances in which allowing or preventing the practice will be consistent with promoting the pursuit of increased economic welfare within the internet ecosystem. The complexities of the interactions taking place across multiple two-sided platforms pose real challenges to the traditional economic modelling used in the past to provide insights into the behaviours in firms and markets. The FCC is not alone in facing these challenges: European regulators are in a similar position in respect of implementing the European Commission's net neutrality Regulation.

However complex and difficult the task may be, regulators and competition law adjudicators have no option but to work their way through the various cases drawn to their attention. The internet ecosystem is fast-moving and constantly evolving, with new technologies, products, services, business models and pricing structures emerging at an unprecedented pace. Zero-rated applications are but one of the many challenges faced. Regulators and adjudicators cannot resile from the task of evaluating the cases brought to their attention.

As a first step to assisting regulators and adjudicators in assessing cases of zero rating brought before them, a set of guidelines identifying the cases more or less likely to be of concern is indicated. BEREC (2016) provides guidance for European regulators, from which it might be possible to develop guidelines to assist in the United States context. However, as we identified in section 1, these focus more on the legal implementation of the European Regulation than on providing guidance to assist in assessing the economic merits of zero-rating cases. That is not to say the guidelines are unhelpful – rather, it is a comment on their limitations.

In this paper, we have sought to evaluate and build upon the BEREC guidelines by focusing primarily on the economic implications of zero-rating. Using a combination of existing net neutrality literature and the application of simple economic models to both the theoretical claims made by net neutrality advocates and some of the complex real-world circumstances where zero-rating has been or may potentially be applied, we have developed a set of questions designed to tease out some of the nuances and trade-offs that the pricing policy invokes. These are:

Question 1. ***What very close or perfectly substitute applications accessible over the ISP's connection, costing the same to deliver, are likely to be foreclosed by the zero-rated application(s)?***

Question 2. ***Does usage of the zero-rated applications actually cost the ISP less than equivalent usage of non-zero-rated applications?***

Question 3. ***Is zero-rated access to a subset of applications primarily intended to increase the number of individuals using the internet?***

Question 4. ***Who has requested that an instance of zero-rating be investigated?***

Question 5. ***Do consumers of the zero-rated application and its rivals make payments to applications providers separate from their payments to ISPs?***

This list of questions is neither complete nor exhaustive. The theoretical and empirical analysis of this topic is still very new. There are many complex and difficult trade-offs to consider, and much still to learn. Theoretical analysis can inform, but ultimately, the applicability of the theoretical models is contingent upon how well they match the complex and uncertain circumstances in the real world. We hope that the questions we have formulated, and the discussion around each of them in this paper will facilitate the making of both good policy and good decisions on the cases of zero-rating that will inevitably arise.

References

Bauer, J., & Obar, J. (2014). Reconciling political and economic goals in the net neutrality debate. *The Information Society 30*, 1-19..

Body of European Regulators for Electronic Communications, (2016). *Draft BEREC Guidelines on implementation by National Regulators of European net neutrality rules.* Retrieved July 27 from file:///C:/Users/howellbr/Downloads/6075-draft-berec-guidelines-on-implementation_0.pdf .

Bring, D., Leighton, W., LoPiccalo, K., Matraves, C.,Schwarz, J., & Waterman, D. (2015). The year in economics at the FCC, 2014-2015. *Review of Industrial Organization 47*(4), 437-62.

Carlton, D., & Perloff, J. (2005). *Modern Industrial Organization.* 4th edition. Boston: Addison & Wesley.

Cho, S., Qui, L., & Bandyopadhyay, S. (2016). Should online providers be allowed to subsidize content? An economic analysis. Working paper available from http://ssrn.com/abstract=2732077 .

Choi, J., Jeon, D-S. & Kim, B-C (2015). Network Neutrality, Business Models, and Internet Interconnection, *American Economic Journal: Microeconomics* 7, 104-141.

- (2014). *Asymmetric Neutrality Regulation and Innovation at the Edges: Fixed vs Mobile Networks*. CESIFO Working Paper No. 4974,

Choi, J., & Kim, B-C. (2010), Net neutrality and investment incentives, *RAND Journal of Economics 41*(3), 446-71.

Economides, N. (2011), Broadband openness rules are fully justified by economic research, *Communications and Strategies 84*, 127-51.

Economides, N. & Hermalin, B. (2015). The Strategic Use of Download Limits by a Monopoly Platform, *Rand Journal of Economics* 46, 297-327.

- (2014).
- (2012). The Economics of Network Neutrality, *Rand Journal of Economics* 43, 602-629.

Economides, N., & Täg., J. (2012). Network Neutrality on the Internet: A Two-Sided Market Analysis, *Information Economics and Policy* 24, 91-104.

Elaluf,-Calderwood, S., & Layton, R. 92015). *Zero –rating: do hard rules protect or harm consumers and Competition? Evidence from Chile, Netherlands & Slovenia*. Paper presented at TPRC, August 2015.

Evans, D., & Schmalansee, R. (2012). *The antitrust analysis of multi-sided platform businesses.* Coase-Sandor Working Paper Series in Law and Economics, 623.

Faulhaber, G. (2011). Economics of net neutrality: A review. *Communications & Convergence Review*, 3, 53–64.

Filistrucchi, L., Geradin, D., van Damme, E., & Affeldt, P. (2014). Market definition in two-sided markets: theory and practice. *Journal of Competition Law and Economics 10*(2), 293-339.

Frieden, R. (2008). A primer on Network Neutrality. *Intereconomics 43*(1), 4-15.

Gans, J. (2015). Weak Versus Strong Net Neutrality, *Journal of Regulatory Economics* 47, 183-200.

Gans, J., & Katz, M.. (2016). Weak versus strong net neutrality: Corrections and extensions, *Journal of Regulatory Economics 50*(1), 99-110. .

- (2016a). *Net Neutrality, pricing instruments and incentives*. NBER Working Paper 22040.

Greenstein, S., Peitz, M. & Valletti, T. (2016). *Net Neutrality: a fast lane to understanding the trade-offs*. NBER Workign Paper 21950.

Guo, H., Hsing, K., & Bandyopadhyay, S. (2013). Broadband network neutrality management and the net neutrality debate. *Production and Operations Management 22*(5): 1287-98.

Hahn, R., & Wallsten, S. (2006). *The economics of net neutrality*. AEI-Brookings Joint Center Working Paper RP05-13.

Howell, B. (2016). *Non-neutral net neutrality: asymmetric regulation strikes again!* TechPolicyDaily.com, April 7 2016. Retrieved August 14 from http://www.techpolicydaily.com/internet/non-neutral-net-neutrality-asymmetric-regulation-strikes-again/.

- (2016a). *The hypocrisy of the anti-paid prioritisation movement.* TechPolicyDaily.com, January 22 2016, Retrieved August 14 2016 from
- http://www.techpolicydaily.com/internet/the-hypocrisy-of-the-anti-paid-prioritization-movement/
- (2014). *Separation anxieties: structural separation and technological diffusion in nascent fibre networks.* Paper presented at the International Telecommunications Society's 20th Biennial Conference, November 30 – December 3, 2014, Rio de Janeiro.
- (2010). *Flat-rate tariffs and competitive entry in Telecommunications Markets.* Wellington: New Zealand Institute for the Study of Competition and Regulation. Retrieved August 14 2016 from
- http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/4053/Tariff_Structure_and_Competitive_Entry_BH.pdf?sequence=1
- (2003). *Building best practice broadband in New Zealand: bringing infrastructure supply and demand together*. Wellington: New Zealand Institute for the Study of Competition and Regulation. Retrieved August 14 2016 from http://researcharchive.vuw.ac.nz/xmlui/bitstream/handle/10063/3887/building_best_practice_060603.pdf?sequence=1 .

Howell, B., & Layton, R. (2016). *Differentiating Neutrality and Non-Discrimination: a Politico-Techno-Economic Evolutionary History of the Internet and the Network Neutrality Debate.* American Enterprise Institute, forthcoming.

Katz. M. (2016). Wither U.S. Net Neutrality Regulation? Technology Policy Institute Working Paper, Retrieved July 23 2016 from https://techpolicyinstitute.org/policy_paper/wither-u-s-net-neutrality-regulation/ .

Krämer, J., Wiewiorra, L., & Weinhardt, C. (2013), Net neutrality: a progress report. *Telecommunications Policy 37*(9), 794-813.

Kourandi, F., Krämer, J. & Valletti, T. (2015). Net Neutrality, Exclusivity Contracts and Internet fragmentation, *Information Systems Research* 26(2), 320-338.

Laffont, J., & Tirole, J. (2001). *Competition in Telecommunications*. Cambridge, Massachusetts: The MIT Press.

Layton, R. (2016). *BEREC's guidelines on zero rating are another "economics-free zone": populism is the guiding principle.* Retrieved July 27 2016 from http://roslynlayton.com/berecs-guidelines-on-zero-rating-are-another-economics-free-zone-populism-is-the-guiding-principle/ .

Lee, R., & Wu, T. (2009). . Subsidizing Creativity through Network Design: Zero-Pricing and Net Neutrality, *Journal of Economics Perspectives* 23, 61-76.

Leiner, B., Cerf, V., Clark, D., Kahn, R., Kleinrock, L., Lynch, D., Postel, J., Roberts, L., & Wolff, S. (2003). *A Brief History of the Internet*. Retrieved June 18 2016 from http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet .

Lemley, M., & Lessig, L. (2000). *The end of End-to-End: preserving the architecture of the internet in the Broadband era*. U.C. Berkeley Public Law and Legal Theory Research Papwer No. 36.

Peitz, M., & Schuett, F. (2015). *Net Neutrality and the Inflation of Traffic.* TILEC Discussion Paper No. 2015-006, Tilburg University.

Schuett, F. (2010). Network neutralitry: a survey of the economic literature. *Review of Network Economics 9*(2), 1-14.

Van Schewick, B. (2016), T-Mobile's Binge On Violates Key Net Neutrality Principles. Mimeo retrieved March 31 2016 from https://cyberlaw.stanford.edu/downloads/vanSchewick-2016-Binge-On-Report.pdf

- (2015), *Network Neutrality and zero-rating*. Attachment to Barbara van Schewick's Ex Parte in the Matter of Protecting and Promoting the Open Internet submitted February 19, 2015 to the Federal Communications Commission GN Dkt. No. 14-28. http://apps.fcc.gov/ecfs/document/view?id=60001031582

Wu, T. (2003). .Network Neutrality, Broadband Discrimination, *Journal of Telecommunications and High Technology Law* 2, 141-178.

Yoo, Christopher S., 2005. Beyond Network Neutrality. *Harvard*

Yoo, C. (2014). Wickard for the Internet? Network Neutrality after *Verizon v. FCC*. *Federal Communications Law Journal 66*(3), 415-466.